

"Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
Program information and without liability
for any use made thereof."

Interim Report

E7.3-11105

CR-135574

ORSER-SSEL Technical Report 19-73

PROCESSING OF REMOTE SENSING DATA

G. J. McMurtry, S. J. Chung, and N. B. Bolling

ERTS Investigation 082
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA
WITHIN THE SUSQUEHANNA RIVER BASIN

Resource Inventory, Land Use, and Pollution

(E73-11105) PROCESSING OF REMOTE SENSING
DATA Interim Report (Pennsylvania State
Univ.) 20 p HC \$3.00 CSCL 14B

N73-33262

63/13 Unclass
01105

Office for Remote Sensing of Earth Resources (ORSER)
Space Science and Engineering Laboratory (SSEL)
Room 219 Electrical Engineering West
The Pennsylvania State University
University Park, Pennsylvania 16802

Principal Investigators:

Dr. George J. McMurtry
Dr. Gary W. Petersen

Date: May 1973

PROCESSING OF REMOTE SENSING DATA

G. J. McMurtry, S. J. Chung, and N. B. Bolling

Through experience with ERTS-1 and aircraft data, ORSER has developed an effectively operational system for processing both imagery and computer compatible tapes. This system is described below.

Processing Imagery

Photointerpretive equipment, study facilities and imagery storage closets are located in the ORSER Laboratory, in Room 218 of the Electrical Engineering West building. Three light tables (including one with high intensity lights) are provided for the study of transparencies and films of various sizes, either as rolls or as individual frames. Film roll holders on one of the tables permit simultaneous viewing of two film strips, each up to 9 inches wide.

A Bausch & Lomb Zoom 70 Stereoscope is available for use with 70 mm C130 and U2 film or for detailed study of ERTS-1 images. Single images may be viewed at enlargements from 10 to 120X using the single lens attachment. When this attachment is replaced by the rhomboid assembly, image pairs in the 70 mm format may be viewed stereographically. The zoom feature of this unit permits viewing at any scale from 10 to 30X with no adjustment of stereo fusion necessary during the continuous change in scale. The Zoom Stereoscope may be used on its small light table base, with two 70 mm film reel holders attached, or it may be used on a bracket above the high-intensity light table. In this latter arrangement, the stereoscope can be moved at will across two 70 mm film strips mounted simultaneously on the light table, and along their length for a distance of 24 inches.

An Old Delft Stereoscope and two mirror stereoscopes (one with a binocular attachment) are provided for the study of 9 X 9 inch stereo

pairs. Two of these stereoscopes permit viewing at two scales: 1.5X and 4X for the Old Delft, and 1X and 6X for the mirror stereoscope. Projection equipment is also available in the laboratory, including a 35 mm slide projector and a Visucom high intensity overhead projector.

A Bausch & Lomb Zoom Transferscope has just been received. This instrument permits projection of opaque or transparent images onto a plain surface or another image, with the capability of magnification in any direction, or selectively in a single direction, from 1 to 14X. With this instrument a photograph can be projected onto a computer-generated character map, with adjustment for the line and element distortion inherent in the high speed printer output.

In addition to the ORSER equipment listed above, a completely equipped photogrammetric and photointerpretation laboratory in the Department of Civil Engineering is available for use by ORSER personnel. Particular pieces of equipment found to be useful in this laboratory include a Kelsh plotter and an American Optical Delineascope which has been used to project 70 mm U2 imagery. An Itek reader-printer for 35 and 70 mm film and a Saltzman Projector for aerial photography are available in the Department of Geology. Various types of equipment for collecting ground truth and producing other related information exist in the several departments of the co-investigators and are available to ORSER.

For image reproduction, The Pennsylvania State University is fortunate to have a wide variety of photographic and reproduction facilities available. Still Photography Services, of the Division of Instructional Services, is staffed and equipped to provide professional, high quality, products meeting all research and publication requirements. Enlargements, contact prints, and slides in color as well as black and white are readily obtained. Complete Diazo and Ozalid facilities are also available on campus.

Processing Digital Data

The ORSER system for processing MSS digital tapes¹ was developed for use by a wide variety of researchers working in remote sensing at The Pennsylvania State University. These potential users represent many disciplines and have a wide range of experience and skill in computer usage.

Computer Processing Facilities

Automatic data processing equipment utilized by ORSER personnel is primarily located at The Pennsylvania State University Computation Center. The principal computer is the IBM System 370 Model 165, consisting of a main frame and attached devices for input and on-line storage. A Remote Job Entry (RJE) system is available and is the most common method used by ORSER to process remote sensing data. The RJE system permits use of IBM 370/165 from any of several remote terminals. Any compatible terminal may be used to process data via the RJE system, including equipment at non-University Park locations tied in by long-distance telephone connections. An IBM 360/50 is also available as well as an IBM 1401 computer. Display equipment at the Computation Center includes an AGT-30 ADAGE graphics terminal and computer, and a CalComp 564 plotter. Extensive analog and hybrid processing equipment, as well as other digital systems, are also available on campus.

The IBM 370/165 is dedicated to general University research and educational uses as well as to similar non-University uses. Users may have access to the computer in any of three ways: (1) central and remote high speed dispatch points operated by the Computation Center, (2) slow speed RJE terminals using IBM 2741 or similar terminals supported by the user or by the Computation Center, and (3) intermediate speed remote batch terminals, such as the IBM 2780, supported by the user or the Computation Center. The processing system for MSS data

¹This system is described by Borden in: "A Digital Processing and Analysis System for Multispectral Scanner and Similar Data," Remote Sensing of Earth Resources, (F. Sharokhia, ed.), Vol. 1, University of Tennessee Space Institute, Tullahoma, Tennessee.

was developed for use with any of these entry points. ORSER investigators have direct access to the RJE terminals and one has been ordered for the ORSER laboratory. These terminals are used for most developmental work. Bulk output for final runs is directed from an RJE terminal to any of the high speed terminal sites. No program card decks need to be input, as the MSS data processing programs are kept in library files. Files for building control information or for storing output are available to the user. MSS data is input from magnetic tapes which, along with user-owned working tapes, are managed by the Computation Center. Non-University users as well as University users may join the system, either locally or via long distance telephone lines.

A standard digital tape format was designed within which all known MSS sources can conveniently be placed. More than one file per tape is allowed as well as a continuation of a file to another tape. Within the file, four kinds of records exist: (1) identification records, (2) table of contents records, (3) MSS response records, and (4) history records. Each MSS response record consists of a complete scan line. Each scan line is numbered and scan lines are always in ascending order in a file. A working file will usually contain one or more small parts of the whole data set. The table of contents is particularly useful in such cases in avoiding costly searching for data which are not present in the file.

The system is couched in a multivariate framework. Although it is understood that some operations do not require this statistical basis, this approach is, overall, most appropriate. Each observation, identifiable by scan line and element number, consists of a vector with as many components as there are channels. At present, each vector is composed of just MSS response values. It is anticipated, however, that the vectors will be augmented by transformed scanner data; or by additional, nonscanner, data such as topographic information.

The system is not in a conversational mode, where the user and the system dynamically interact during processing. Each program accepts input control specifications, processes the MSS data according to the specifications, and outputs the results. The user prepares the control

specifications for each program. Although the system is non-conversational, the preparation of the control specifications by a user operating from an RJE terminal is conversational. For non-RJE operation, control specifications are made and entered into the system by punched cards. All control specifications on the RJE are identical in format to the corresponding punched cards.

The Digital Data Processing System

The programs discussed here are all operational and are documented at the user's level. Although many other programs are used, those discussed here illustrate the general approach to the processing of MSS tapes. Detailed descriptions of ORSER programs currently available may be found in ORSER/SSEL Technical Report 10-73.

The digital tape processing system for MSS data described here is regularly run for production and has been extended to meet the needs of various related projects. The system was designed to be easily augmented, typified by the addition of a number of supervised and unsupervised analysis and classification algorithms. The general procedure to be employed for a previously unstudied area or type of target will be presented and illustrated here. The procedure to be employed for areas or targets which have been previously investigated differs slightly from that shown here and is less complex.

The first step is to select the particular targets and areas of interest, primarily using maps. Consultation of the catalogues of imagery and digital tapes will indicate what data are available and their quality. Tapes corresponding to the selected scenes are chosen and the areas of useful data are specified. Subsets are then produced on separate tapes, using the SUBSET program¹. These subsets are prepared to gain rapid processing and short turn around time. It is likely that this step has already been done in the process of cataloguing and storing ERTS tapes by ORSER, in which case the appropriate library subset tapes would be selected directly.

¹Complete program descriptions may be found in ORSER-SSEL Technical Report 10-73.

A run is then made with the NMAP program to show the overall pattern of the data. This program is written to map element brightness, using all channels or any subset of channels. The norm of each multivariate vector is taken as the measure of brightness. The norm is then converted to a percentage of the maximum possible value. This value is translated to the mapping symbol for the percentage range within which it falls. The process is repeated for every element in every scan line in the data blocks specified by the user. Output from the NMAP program consists, then, of a brightness map.

The UMAP program is run next, to map areas of local spectral uniformity. Each element is compared with its near neighbors using the euclidean distance between spectral signatures as the measure of similarity or dissimilarity. If the largest distance is smaller than a value specified by the user, then the symbol for uniformity is assigned to that element. One or more categories of uniformity can be mapped according to distances specified by the user. All elements with distances from their neighbors greater than those specified are mapped as contrasts. The map output shows the pattern of uniformity and contrasts from which the user can designate coordinates for training areas for the targets of interest and determine high contrast boundaries between uniform areas.

Signatures and associated statistics are next obtained by the use of the STATS program, which computes the multivariate statistics for one or more training areas obtained from UMAP or similar output. The user designates a training area by line and element coordinates and the program computes the statistics for all of the data which fall within the boundaries. The mean and standard deviation vectors are found, and the correlation and variance-covariance matrices are computed as well as the eigenvalues and eigenvectors of these matrices. Frequency histograms for selected channels are also computed.

When most of the targets have been identified by training areas, a classification run is made using the classifier or classifiers deemed most appropriate for the mix of targets under consideration. A variety of classification programs are available, including parametric and

non-parametric classifiers with either linear or quadratic discriminant functions. Preprocessing before classification is also possible, using programs for normalization, principal components, etc. The output of these programs is in the form of a character (or digital) map, with each category of classification represented by a unique symbol.

Digital maps are useful primarily as working maps for the user in the analysis of MSS data. They are inherently distorted in the length-to-width relationship because of the fixed number of lines and characters per inch of high-speed printer output. The LMAP program, yielding output on the CalComp plotter, is intended for the production of distortion-free, finished copy, line maps. There are three main advantages to line maps when compared to character maps: (1) orthographic maps to a selected scale can be made, (2) photographic overlays can be prepared for these maps (this is quite important in the comparison of classification results with corresponding imagery), and (3) legible maps for publication purposes can be prepared.

An example of the use of the programs described above is given in Figures 1 through 11. The MSS data used for this analysis came from ERTS-1 scene 1028-15295, scanned on August 20, 1973. This is an area northeast of Clearfield, Pennsylvania, on which U.S. Route 80 and the West Branch of the Susquehanna River cross. The location of the test site is shown in Figure 1, which was taken from two 7 1/2 minute USGS quadrangle maps. As these maps were printed in 1959, before Route 80 was constructed, the highway has been drawn in on the figure. Figure 2 and 3 show map output for NMAP and UMAP, respectively. The strip of low brightness in Figure 2 follows the river, as does the blank (non-uniform) area shown in Figure 3. Basic statistics for the "stripmine" category, obtained by the STATS program are shown in Figure 4. Statistics from a series of such sample sites are input to a classification program. Figure 5 shows the output from the DCLASS, program which classifies according to a minimum euclidean distance algorithm. Only two categories are represented by symbols; unclassified elements are left blank. LMAP output using data from the DCLASS program is shown in Figure 6.

It frequently happens that a sample target is not of sufficient size or area to lend itself to categorization using the STATS program.

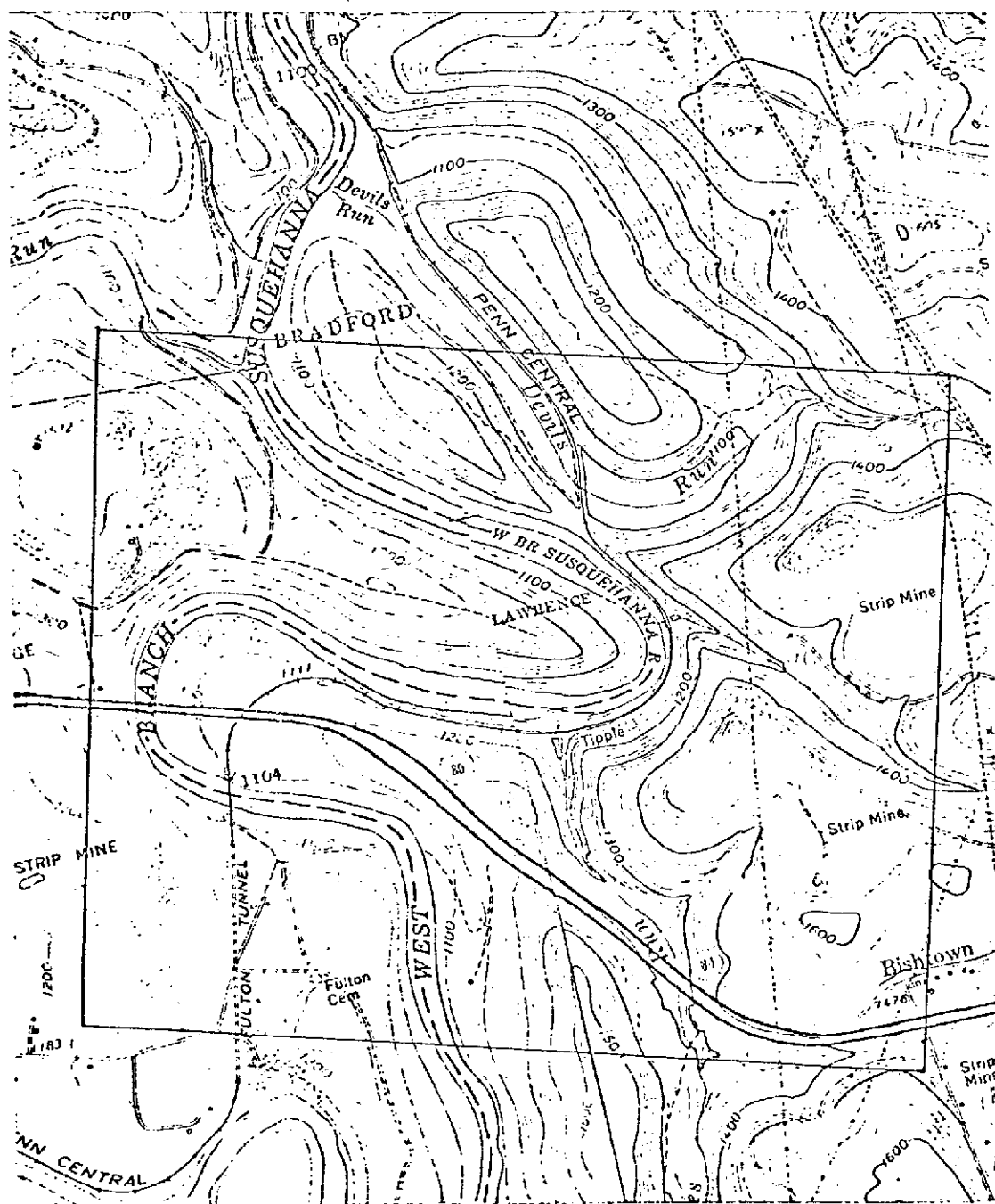


Figure 1: Test site northeast of Clearfield, Pennsylvania. (Taken from USGS 7 1/2 minute quadrangle maps, "Clearfield" and "Leontes Mills," both printed in 1959.)

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
 ENDING LINE 1059
 BEGINNING ELEMENT 2470
 ENDING ELEMENT 2525
 LINE INCREMENT 1
 ELEMENT INCREMENT 1

	2470	2475	2480	2485	2490	2495	2500	2505	2510	2515	2520	2525
1030	XXXXXXXXXXXXX-	-XXXX-XXXXXXXXXX	-XXXXXXXXXXXXX	----	XXXXXXXXXX							
1031	XXXXXXXX-XXX-X-	XXXXXXXXXXXXXX	-XXXXXXXXXXXXX	----	XXXXXXXXXX							
1032	XXXXXXXX-XXXXXX-	-XXXXXXXXXXXXX	-XXXXXXXXXXXXX	----	XXXXXXXXXX							
1033	XXXXXXXX-XXXXXX-	-XXXXXXXXXXXXX	-XXXXXXXXXXXXX	----	XXXXXXXXXX							
1034	XXXXXXXX-XXXXXX-	----	-XXXXXX-	-X-	----	-XX-	-XX					
1035	XXXXXXXXXXXXXXXXXX	-XXXXX-	----	-X-	----	-X-	-XXX					
1036	X-XX-XXXXXXXXXXXXXXXXXX	----	----	----	----	----	XX-	-XXXX				
1037	X-XX-XXXXXXXXXXXXXXXXXX	-XXX-	----	XXXXXX-XX-X-	-XX-X-							
1038	X-XXXXX-XX-XX-XXXXXX	-XXXXXX-XXXXXX-XX-	-XXXXXX-XXXXXX-XX-									
1039	XX-XX-XXXXXX-XXXXXX	-XXXXXX-XXXXXX-XXXXXX	-XXXXXX-XXXXXX-XXXXXX									
1040	XXXXXXXX-XXXXXX-X-	----	XXXXXXXXXXXXXXXXXX	----	XXXXXXXXXX	----						
1041	XXXXXXXX-XX-XXXXXX-	----	XXXXXXXXXXXXXXXXXX	-XXX-	-XXXXXX-XXXXXX-							
1042	XXXXXXXX-XXXXXX-XXXX-	----	XXX-	-XXXXXXXXXXXXXXXXX-								
1043	XXXXXXXX-XXXXXXXXXXXXX	----	-	----	XXXXXXXXXXXXX-XX							
1044	XXXXXX-XXX-XXXXXXXXXXXXX	----	----	XXXXXXXXXXXXXXXXXX-								
1045	XXXXXXXX-XX-XXXXXXXXXXXXXXXXXX	----	XXXXXX-XXXXXXXXXXXXX-									
1046	XXXXXXXXXX-	-XXXXXXXXXXXXXXXXXXXXX	----	XXXXXXXXXX-XXXXXX								
1047	XXXXXXXX-XXX-	----	XXXXXXXXXXXXX-XXXXXXXXXXXXX	----								
1048	XXXXXXXX-XXXX-XXXXXX-	-XXXXXXXXXXXXX-	XXXXXXXXXXXXX-XXXXXX	----	XX-							
1049	-----XXXXXXXXXXXXXX	XXXXXXXXXXXXX-XXXXXXXXXXXXX	XXXXXX									
1050	-----XXXXXXXXXXXXXX	-XXXXXXXXXXXXX-XXXXXX										
1051	XXXXX-XX-XXXXXXXXXXXXXX	-XXXX-XXXXXXXXXXXXX-	XXXXXX-XXXXXX									
1052	XXXXX-XX-XXXXXXXXXXXXXX	-X-X-XXXXXX-XXXXXX-XXXXXX										
1053	XXXX-XX-XXXXXX-XXXXXX-XXXXXX											
1054	XXXXX-XX-XXXXXX-XXXXXX-XXXXXX											
1055	XXXXX-XX-XXXXXX-XXXXXX-XXXXXX											
1056	XX-XX-XXXXXX-XXXXXX-XXXXXX											
1057	XXXXXXXX-XXXXXX-XXXXXX-XXXXXX-XXXXXX											
1058	XXXXXXXXXXXXXXXXXXXXX-XXXX-	-XXXXXX-XXXXXX-XXXXXX	----	XX								
1059	XXXXXXXXXXXXXXXXXXXXX-XXXXXX	-XXXXXXXXXXXXX-XXXXXX-XX-XXX										

Figure 2: Brightness map (NMAP).

CHANNELS USED : 1 2 3 4

MEANS AND STANDARD DEVIATIONS FOR GIVEN CHANNELS

29.55	26.86	30.78	13.83
3.29	4.80	5.21	3.79

VARIANCE-COVARIANCE MATRIX

10.82			
14.91	23.09		
0.78	0.23	27.16	
-3.27	-5.68	18.00	14.33

CORRELATION MATRIX FOR GIVEN CHANNELS

1.00			
0.94	1.00		
0.05	0.01	1.00	
-0.26	-0.31	0.91	1.00

EIGENVALUES COMPUTED FROM CORRELATION MATRIX.

EIGENVALUES WITH THEIR ASSOCIATED PERCENTAGES:

2.21	1.70	0.06	0.03
55.4	42.4	1.4	0.8

EIGENVECTORS:

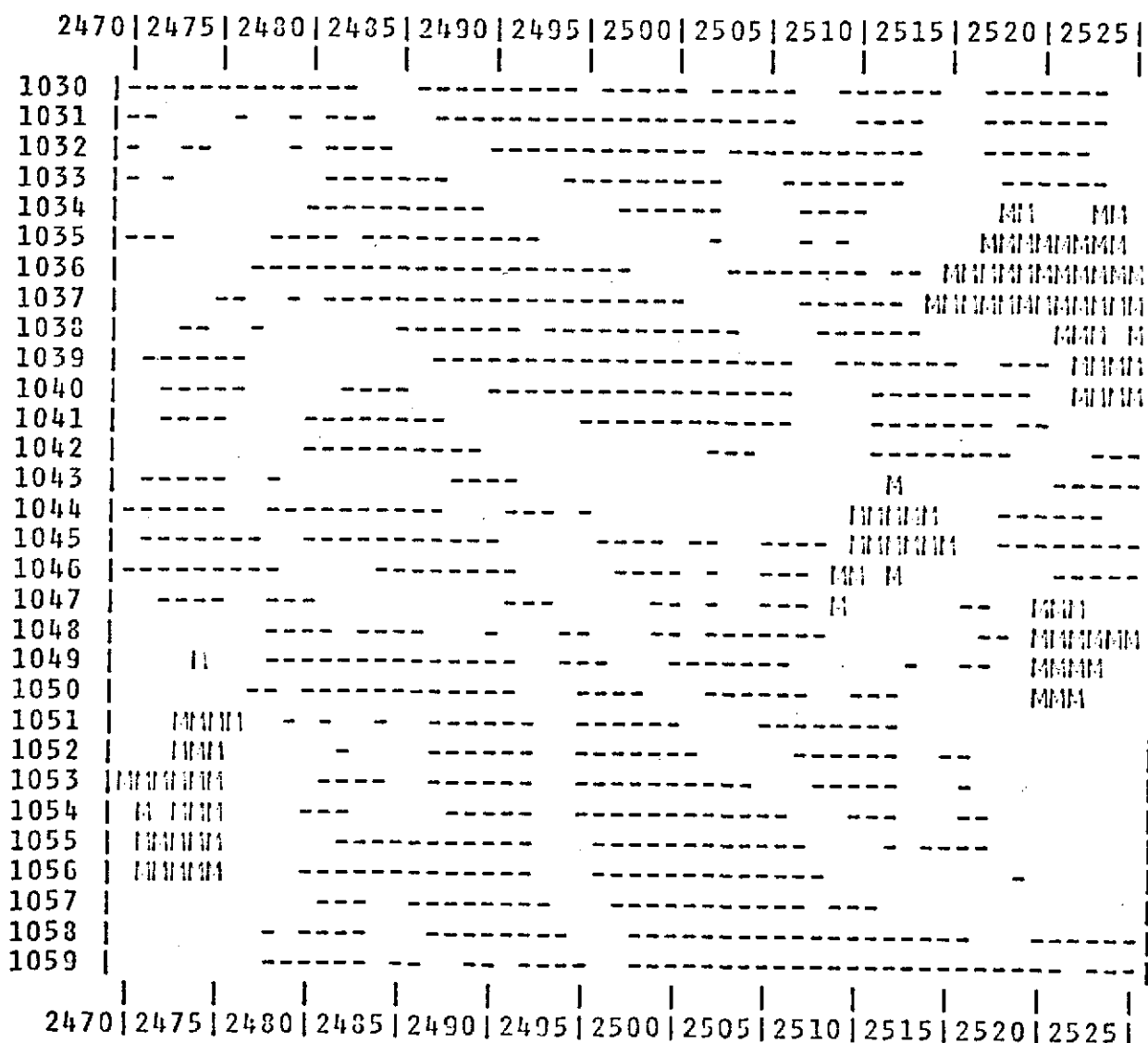
0.5227	-0.4645	-0.7059	-0.1129
0.5416	-0.4368	0.6347	0.3362
-0.3800	-0.6257	0.2327	-0.6402
-0.5376	-0.4493	-0.2115	0.6814
DET.=	0.705D-02		

SAMPLE CATEGORY : STRIPMINE

Figure 4: Statistics of sample areas for a category obtained by STATS.

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
 ENDING LINE 1059
 BEGINNING ELEMENT 2470
 ENDING ELEMENT 2525
 LINE INCREMENT 1
 ELEMENT INCREMENT 1



SYMBOL M : STRIPHINE

SYMBOL - : VEGETATION

Figure 5: Classification map using signatures obtained by STATS.

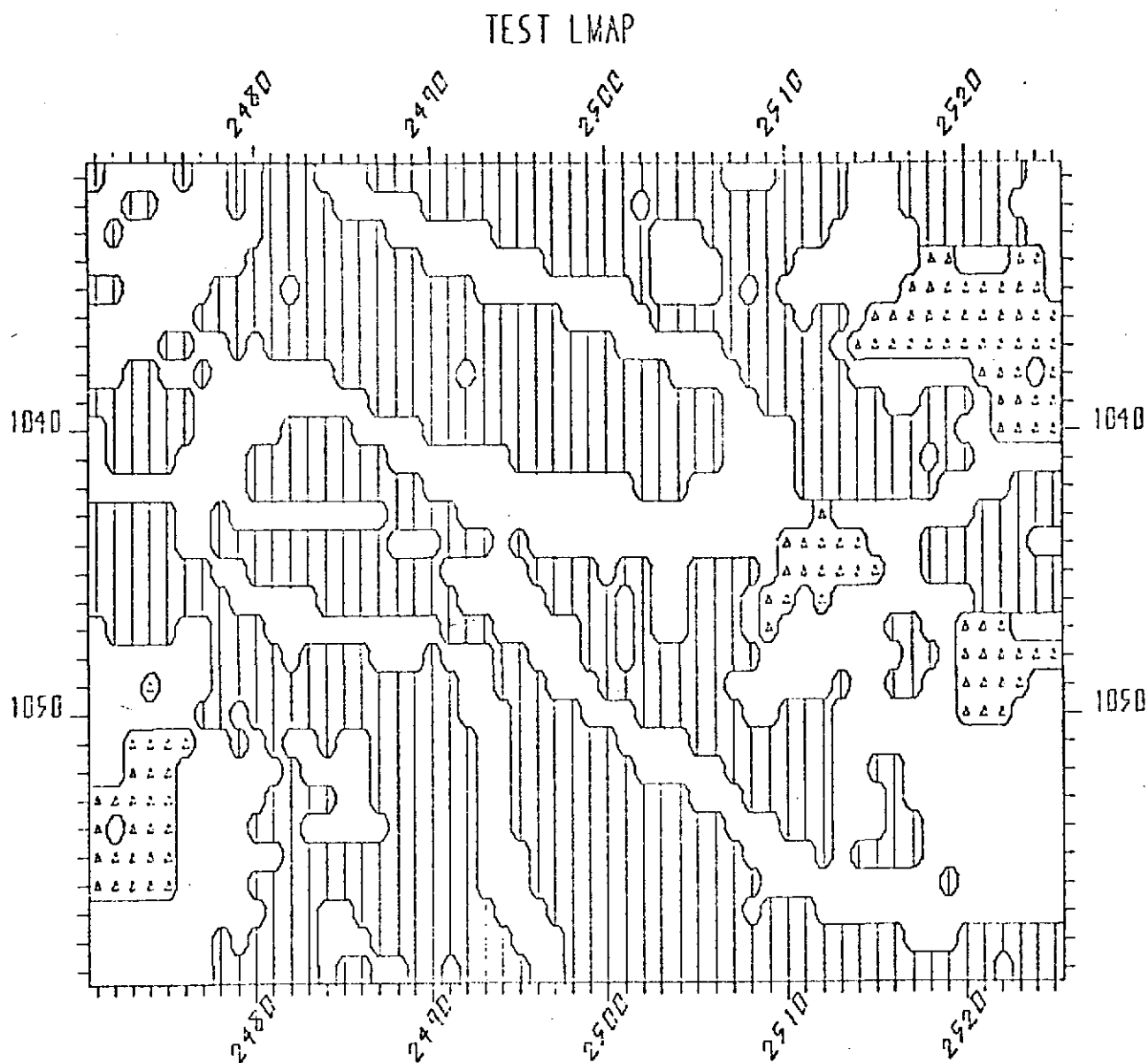


Figure 6: Preliminary classification map of the Clearfield area (LMAP).

Such targets may be linear features such as streams, or a series of small scattered features which are not large enough to be represented as uniform areas by UMAP. In such cases, these areas are defined for analysis by an unsupervised classifier which develops its own set of spectral signatures and statistics using a clustering algorithm. The map output of one such program, DCLUS, is shown in Figure 7. The accompanying statistics are shown in Figure 8, and the input data for DCLASS for the categories to be classified are shown in Figure 9. A comparison of Figure 7, with the DCLASS output of Figure 5 reveals that DCLUS was able to map some features which could not be mapped by DCLASS with STATS signatures. Map output from DCLASS, using DCLUS signatures, is shown in Figure 10. Figure 11 shows this same classification in LMAP form.

The approach employed for change detection or where a temporal dimension is involved is similar to the approach for non-temporal analyses in many respects. The major difference is in the establishment of permanent training areas for supervised analysis and classification and permanent analysis areas for unsupervised analysis and classification. These areas must be selected and specified more carefully and with more refinement than when the temporal dimension is not of interest.

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
 ENDING LINE 1059
 BEGINNING ELEMENT 2470
 ENDING ELEMENT 2525
 LINE INCREMENT 1
 ELEMENT INCREMENT 1

	2470	2475	2480	2485	2490	2495	2500	2505	2510	2515	2520	2525
1030												
1031												
1032												
1033												
1034												
1035												
1036												
1037												
1038												
1039												
1040												
1041												
1042												
1043												
1044												
1045												
1046												
1047												
1048												
1049												
1050												
1051												
1052												
1053												
1054												
1055												
1056												
1057												
1058												
1059												

CHANNELS USED : 1 2 3 4
 INITIAL CRITICAL DISTANCE : 8.5

Figure 7: Signature map by DCLUS.

UN-NORMALIZED CATEGORY SPECIFICATIONS

CHANNELS-	1	2	3	4
1 -	19.74	12.03	41.97	26.84
2 =	29.33	26.80	28.50	12.22
3 +	26.85	22.80	41.38	22.61
4 X	21.84	14.02	49.40	31.55
5 *	20.37	14.79	21.24	9.87
6 .	20.41	14.24	30.24	17.24
7 M	35.67	37.00	36.67	15.00
8 S	32.50	32.00	49.50	26.50

DISTANCES OF SEPARATION FOR CATEGORIES

	1 -	2 =	3 +	4 X	5 *	6 .	7 M	8 S
1 -	0.0	26.6	13.6	9.3	26.9	15.3	32.3	24.9
2 =	26.6	0.0	17.2	32.1	16.8	16.3	14.8	26.1
3 +	13.6	17.2	0.0	15.7	26.0	16.4	19.0	14.1
4 X	9.3	32.1	15.7	0.0	35.6	24.0	34.0	21.5
5 *	26.9	16.8	26.0	35.6	0.0	11.7	31.5	39.0
6 .	15.3	16.3	16.4	24.0	11.7	0.0	28.2	30.3
7 M	32.3	14.8	19.0	34.0	31.5	28.2	0.0	18.2
8 S	24.9	26.1	14.1	21.5	39.0	30.3	18.2	0.0

CHANNELS USED : 1 2 3 4
INITIAL CRITICAL DISTANCE : 8.5
SAMPLE SIZE : 500

Figure 8 : Statistics obtained by DCLUS.

CHANNELS USED: 1 2 3 4

CATEGORY NAME	NUMBER	SYMBOL	LIMIT
STRIPMINE1	1	M	8.5
STRIPMINE2	2	M	8.5
PAVEMENT	3	>	8.5
RIVER	4	R	8.5
VEGETATION1	5	-	8.5
VEGETATION2	6	-	8.5
VEGETATION3	7	-	8.5

UN-NORMALIZED CATEGORY SPECIFICATIONS

CHANNELS-	1	2	3	4
1 M	35.67	37.00	36.67	15.00
2 M	29.33	26.80	28.50	12.22
3 >	26.85	22.80	41.38	22.61
4 R	20.37	14.79	21.24	9.87
5 -	19.74	12.03	41.97	26.84
6 -	21.84	14.02	49.40	31.55
7 -	20.41	14.24	30.24	17.24

DISTANCES OF SEPARATION FOR CATEGORIES

	1 M	2 M	3 >	4 R	5 -	6 -	7 -
1 M	0.0	14.8	19.0	31.5	32.3	34.0	28.2
2 M	14.8	0.0	17.2	16.8	26.6	32.1	16.3
3 >	19.0	17.2	0.0	26.0	13.6	15.7	16.4
4 R	31.5	16.8	26.0	0.0	26.9	35.6	11.6
5 -	32.3	26.6	13.6	26.9	0.0	9.3	15.3
6 -	34.0	32.1	15.7	35.6	9.3	0.0	24.0
7 -	28.2	16.3	16.4	11.6	15.3	24.0	0.0

Figure 9: Category specifications and separations for DCLASS.

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
ENDING LINE 1059
BEGINNING ELEMENT 2470
ENDING ELEMENT 2525
LINE INCREMENT 1
ELEMENT INCREMENT 1

	2470	2475	2480	2485	2490	2495	2500	2505	2510	2515	2520	2525
1030												
1031												
1032												
1033												
1034												
1035												
1036												
1037												
1038												
1039												
1040												
1041												
1042												
1043												
1044												
1045												
1046												
1047												
1048												
1049												
1050												
1051												
1052												
1053												
1054												
1055												
1056												
1057												
1058												
1059												

SYMBOL M : STRIPMINE
SYMBOL R : RIVER

SYMBOL > : HIGHWAY
SYMBOL - : VEGETATION

Figure 10: Classification map using signatures obtained by DCLUS.

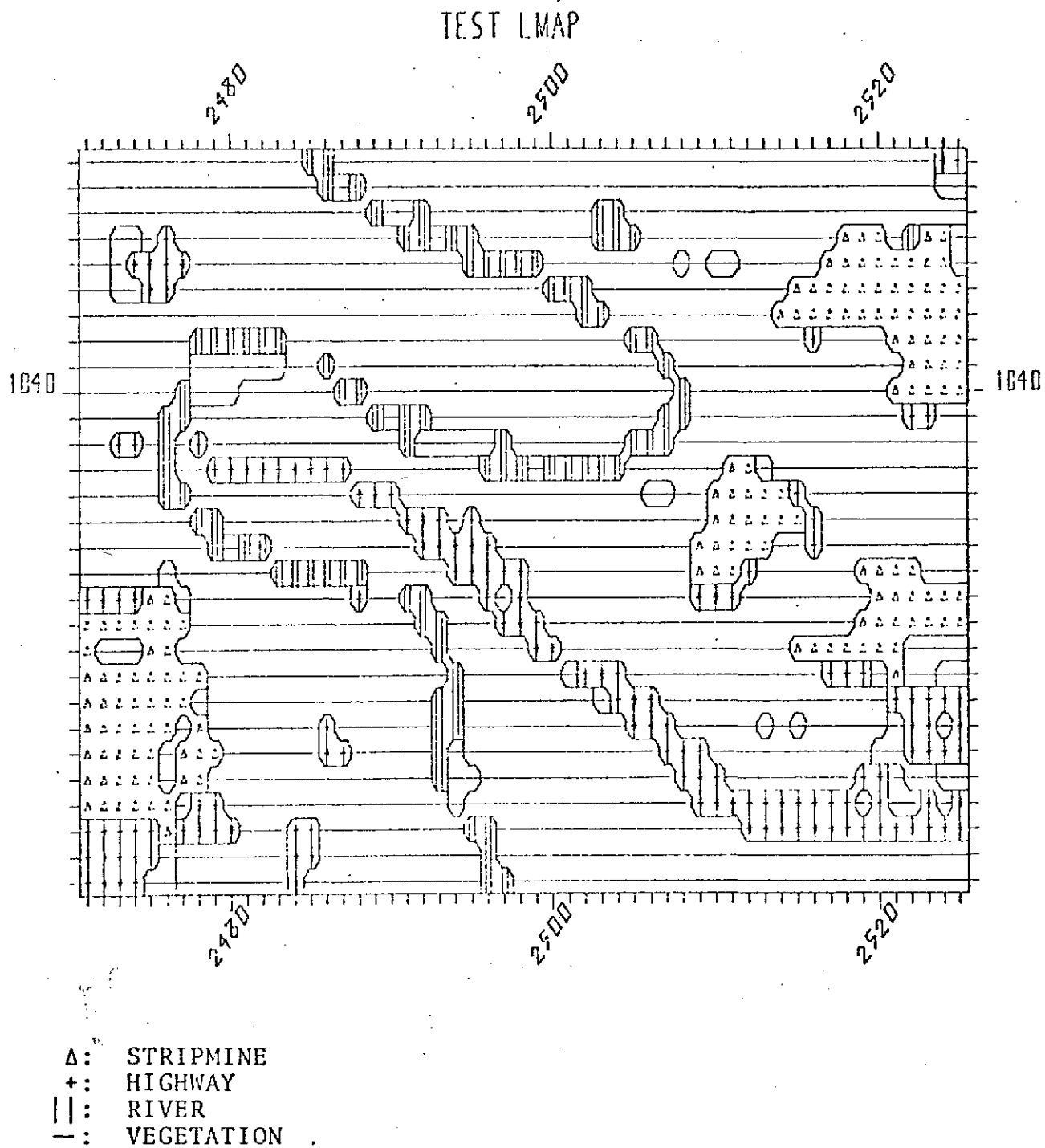


Figure 11: Classification map of the Clearfield area (LMAP).